



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2009

A new look at radiosonde data prior to 1958

Grant, Andrea N ; Brönnimann, Stefan ; Ewen, Tracy ; Nagurny, Andrey

Abstract: Historical radiosonde data are known to suffer from inhomogeneities. The first radiosonde intercomparison was made at Payerne, Switzerland, in 1954, and a major international effort to standardize the network, including launch times, was made for the International Geophysical Year (IGY) in 1957–58. Data from before this period, in some cases extending back as far as 1934, have been viewed with even more suspicion than recent data. These early data are scattered among numerous archives with a variety of station identifier schemes and quality-control procedures, and some of the data have only recently been digitized from paper records. Here, the first systematic compilation of pre-IGY data is made, and a novel quality-assessment technique is applied, which reveals that much of the early data have uncorrected radiation and lag errors, especially in the former Soviet Union. Incorrect geopotential height units and problematic time stamps were also found. The authors propose corrections and present corrected hemispheric fields that show large changes and improved internal consistency in height and temperature across Eurasia compared with uncorrected data. The corrections are important, especially as they have a clear spatial structure that interferes with the planetary wave structure. These corrected data are useful for climate studies and considerably enhance the length and quality of the upper-air record but may not be suitable for trend analysis. Assimilation of the uncorrected data has led to a widespread warm bias in NCEP–NCAR reanalysis in the 1950s.

DOI: <https://doi.org/10.1175/2008JCLI2539.1>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-77127>

Journal Article

Published Version

Originally published at:

Grant, Andrea N; Brönnimann, Stefan; Ewen, Tracy; Nagurny, Andrey (2009). A new look at radiosonde data prior to 1958. *Journal of Climate*, 22(12):3232-3247.

DOI: <https://doi.org/10.1175/2008JCLI2539.1>

A New Look at Radiosonde Data prior to 1958

ANDREA N. GRANT, STEFAN BRÖNNIMANN, AND TRACY EWEN

Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

ANDREY NAGURNY

Arctic and Antarctic Research Institute, St. Petersburg, Russia

(Manuscript received 19 March 2008, in final form 14 November 2008)

ABSTRACT

Historical radiosonde data are known to suffer from inhomogeneities. The first radiosonde intercomparison was made at Payerne, Switzerland, in 1954, and a major international effort to standardize the network, including launch times, was made for the International Geophysical Year (IGY) in 1957–58. Data from before this period, in some cases extending back as far as 1934, have been viewed with even more suspicion than recent data. These early data are scattered among numerous archives with a variety of station identifier schemes and quality-control procedures, and some of the data have only recently been digitized from paper records. Here, the first systematic compilation of pre-IGY data is made, and a novel quality-assessment technique is applied, which reveals that much of the early data have uncorrected radiation and lag errors, especially in the former Soviet Union. Incorrect geopotential height units and problematic time stamps were also found. The authors propose corrections and present corrected hemispheric fields that show large changes and improved internal consistency in height and temperature across Eurasia compared with uncorrected data. The corrections are important, especially as they have a clear spatial structure that interferes with the planetary wave structure. These corrected data are useful for climate studies and considerably enhance the length and quality of the upper-air record but may not be suitable for trend analysis. Assimilation of the uncorrected data has led to a widespread warm bias in NCEP–NCAR reanalysis in the 1950s.

1. Introduction

Radiosonde data are the primary tool by which we understand the vertical structure and circulation of the free atmosphere, both for forecasting purposes and climate studies. Unfortunately, this data record is fraught with problems. Although the apparently divergent temperature trends seen in the radiosonde and satellite records have, after over 15 years of debate, finally been harmonized (Sherwood et al. 2005; Santer et al. 2005; Mears and Wentz 2005; Fu et al. 2004), many questions remain about the homogeneity of the radiosonde record (e.g., Lanzante et al. 2003; Free et al. 2002; Eskridge et al. 2003), limiting confidence in trend analysis.

Several efforts have been made to homogenize the radiosonde record (e.g., Free et al. 2005; Lanzante et al.

2003; Haimberger 2007; Parker et al. 1997; Thorne et al. 2005). Homogenizing surface meteorological data generally relies on comparison with a suitable reference series. Such a reference is normally not available for upper-air data, leading to a variety of techniques that can be broadly classified as statistical [e.g., stationarity of a time series (Lanzante et al. 2003)] or use of an unorthodox reference series [comparison with satellite data (Parker et al. 1997) or highly correlated distant stations (Thorne et al. 2005)]. Intercomparison of different techniques does not routinely reveal the same breakpoints (Free et al. 2002), underscoring the difficulty of the problem. Most attempts extend back only to 1958, the end of the International Geophysical Year (IGY) during which there was a dramatic expansion of the worldwide radiosonde network as well as a standardization of launch times to “00Z and 12Z.” Other homogenization efforts cover only the satellite period (back to 1979) (Parker et al. 1997). In most cases, pre-IGY data were simply too problematic or scarce to bother with, partly owing to the irregular launch times.

Corresponding author address: Andrea Grant, Institute for Atmospheric and Climate Science, ETH Zurich, Universitätsstrasse 16, 8092 Zurich, Switzerland.
E-mail: dr.andrea.grant@gmail.com

Lanzante et al. (2003) began by including data back to 1948 but in a number of cases were forced to discard the pre-IGY data.

Up to now, the pre-IGY data have never even been systematically compiled. The data are scattered among numerous archives, cataloged via multiple station identifier schemes, and have been subjected to different quality control and data culling procedures. It was our hypothesis that some of the discarded earlier data may be usable after quality assessment and correction; therefore, we attempted to compile a comprehensive archive of radiosonde data up to 1957. Much of the work involved detangling the overlapping archives to create a single dataset. These data are a valuable resource for studies of interannual climate variability, case studies, and as input to upcoming reanalysis and reconstruction projects. Although trend studies based on the earlier data are problematic, having a better understanding of the errors and reliability of the earlier data can assist in decisions about the appropriateness of calculating trends.

This study applies a novel quality assessment technique to this new collection of pre-IGY data. The quality assessment process was developed during a project to digitize historical radiosonde data prior to 1948 (Brönnimann 2003b). Monthly mean station data are compared to a reference series that is statistically reconstructed from surface data using relationships derived from National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (NNR) (Kistler et al. 2001). Time series of anomalies from the reference can then be inspected for breakpoints or variance that exceeds prespecified targets, and vertical profiles of the mean anomalies can be inspected for characteristic shapes of common errors. In most cases, the corrections are physically based and are made using independent information; this differs from many homogenization techniques that apply a statistical correction. Section 2 describes the data that were collected and assessed. Section 3 describes the quality assessment technique, with some sample stations being shown in section 4. Overall data quality and findings are discussed in section 5, including before and after figures of Northern Hemisphere station data and a brief discussion of widespread NNR biases due to assimilation of uncorrected data. Finally, in section 6, a conclusion summarizes the results.

2. Data

Data from approximately 1500 radiosonde stations were collected from several archives: 1) the Integrated

Global Radiosonde Archive (IGRA) (Durre et al. 2006) at the National Climatic Data Center (NCDC); 2) the U. S. Air Force Environmental Technical Applications Center tape deck 54 dataset (TD54) at NCAR; 3) the NCDC tape deck 6201 compilation (TD-6201); 4) the Comprehensive Aerological Reference Dataset (CARDS) tape deck 542 archive (CARDS542) covering 1946 and 1947 (Eskridge et al. 1995) from NCAR; 5) data digitized internally within the working group of the authors at ETH Zurich (Brönnimann 2003b; Brönnimann et al. 2005; Ewen et al. 2008); 6) data from the Arctic region of the former Soviet Union, which were recently digitized at the Arctic and Antarctic Research Institute (AARI); and 7) several miscellaneous sources such as Lindenberg, Germany, and Payerne, Switzerland. Significant overlap existed among the archives, complicated by the fact that some were numbered according to the Weather Bureau Army Navy (WBAN) system and others according to the World Meteorological Organization (WMO) scheme: It was not initially clear which stations overlapped. Inconsistent station coordinates presented a further roadblock: for example, pairs of stations with identical WMO numbers but different coordinates were found, yet on further inspection they contained identical data. Other combinations also arose, where pairs of stations had different WMO numbers and same or different coordinates, with identical data in some cases and not in others. Different levels and timestamps were also present in soundings that were otherwise duplicates of each other.

After identifying possible duplicates (identical station numbers or coordinates within 1° latitude and longitude), data were manually compared to identify duplicates. Eight hundred, seventy-nine unique stations were identified, shown in Fig. 1. Because of systematic level deletion, duplicated timestamps, and overall data availability before 1958, sources were prioritized as follows (with records being completed with lower priority sources if additional time periods were available): 1) internally/AARI digitized, 2) TD54, 3) TD-6201, 4) CARDS, and 5) IGRA. IGRA was given a lower priority in the compilation process owing to deletion of certain levels, deletion of early parts of some records, and duplicated timestamps at some stations. The stations were coded by radiosonde type so that the appropriate radiation and lag correction, if necessary, could be applied. Western European stations with no metadata were presumed to have used a Vaisala sonde and Chinese stations were presumed to have used the Soviet Molchanov sonde. A fully cross-referenced station list including all verdicts and corrections is available along with the archive of monthly mean data.

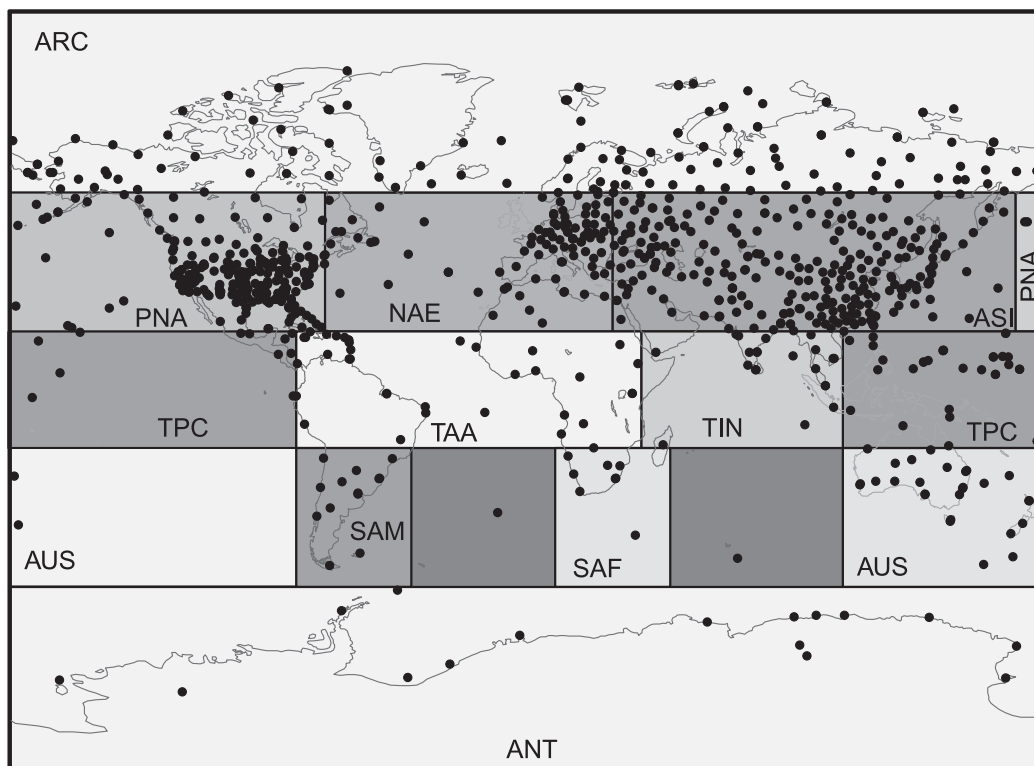


FIG. 1. Map of radiosonde stations used in this study. Shaded areas delineate the regions used for reconstruction input.

3. Quality assessment

Data transmitted on the Global Telecommunications System (GTS) undergo routine quality control measures before being archived at NCDC (Durre et al. 2006; Eskridge et al. 1995; Kalnay et al. 1996; Kistler et al. 2001). Nevertheless, the record, especially in the earlier years, remains problematic (e.g., Lanzante et al. 2003). Having successfully assessed the quality of upper-air data from 1939 to 1944 using a new statistical technique (Brönnimann 2003b), we apply this method to the newly compiled pre-IGY data. The full details of the technique can be found in Brönnimann (2003a). An overview of the technique is presented next.

a. Context and limitations

One of the main problems in homogenizing radiosonde data is a lack of neighboring stations, unlike surface data where suitable stations can normally be found. Various alternatives have been employed, such as self-homogenization techniques (Lanzante et al. 2003), comparison with satellite data (Parker et al. 1997), the use of “buddy” stations showing a high correlation with the station of interest (Thorne et al. 2005), and background fields from ERA-40 reanalysis (Haimberger 2007).

In our approach, a statistically reconstructed monthly mean reference series was generated for both temperature and geopotential height at each station on a subset of pressure levels. Because of the potential errors involved in such a reconstruction, the quality assessment is limited to a small number of physically based errors for which all of the data for a given station (or large, contiguous temporal subsets, if necessary) can be corrected based on a single parameter. This is in contrast to the more statistical homogeneity approaches, which identify errors in individual variables and heights and then adjust small sections of the time series independently of other levels and variables. Our approach limits the number of errors that can be addressed, but even with this conservative approach a large number of errors can be identified and corrected. Our approach also simplifies the identification of errors since we have an expectation of what the possible errors could be based on physical principles, knowledge about operational processing of the early data, and previous work with early upper-air data (Brönnimann 2003b; Ewen et al. 2008).

The data were tested using classical statistical breakpoint detection tests (Alexandersson and Moberg 1997; Lanzante 1996), and we then take a conservative approach

in accepting these breakpoints only when there is meta-data supporting the adjustment. Prior to the IGY, such metadata could include the following: 1) the launch time change (to 0000 and 1200 UTC) when it was likely that other operational changes were made; 2) the IGY itself, also a time of possible operational changes; or 3) a situation in which the change under consideration also occurred at other stations in the same country or network at the same time (e.g., the January 1950 change in reported geopotential height units across the entire Soviet Union). In practice this means that corrections apply for the entire pre-IGY time series or for large and contiguous subperiods (e.g., units correction from beginning of the record to 31 December 1949). This conservative approach relies on an assumption of simplicity regarding operational processing; that is, a station did not arbitrarily jump back and forth between processing routines from one sounding to the next. In some cases an endpoint was not identified: corrections were required up to the IGY, but owing to the limited nature of this study, a specific endpoint beyond this cannot be specified. In summary, the errors that we identify are assumed to stem from operational or instrument calibration errors that are relatively constant for a period of time and then change, leading to corrections that apply to the entire time period under study or large and contiguous subperiods. The outcome improves spatial homogeneity in the early time period. Although the temporal homogeneity is also improved, it may be the case that the corrected data are not suitable for trend analysis.

b. Reconstruction of the reference series

The statistically reconstructed reference series (hereafter “reference” or “reference series”) was generated for each candidate series, that is, for each variable (temperature and height) at five pressure levels (850, 700, 500, 300, and 200 hPa) for each station. The reconstruction is based on a multiple linear regression of upper-air series (NNR interpolated to the station coordinates) as a function of surface data (temperature and pressure). The NNR series were regressed onto the surface predictors for the period from 1960 to 2000 (the model calibration period) to create the statistical model. This model was then applied to the surface predictors in the pre-IGY period to reconstruct a reference series for each variable, height, and station. The model was validated for the period 1948–59 and estimates of skill (see below) were derived during this validation period. Note that the quality of NNR is presumably worse in the early years. This does not affect our reconstruction but tends to lower the skill. Hence, the skill measure is conservative.

The surface predictors were defined for 11 regions (Fig. 1) and encompass the first 15 principal component time series of surface temperatures in the form of standardized monthly anomalies and the first 10 principal component time series of sea level pressure anomalies, for which anomalies were defined based on the 1961–90 mean seasonal cycles. Sea level pressure data were taken from the second Hadley Centre Sea Level Pressure dataset (HadSLP2) (Allan and Ansell 2006) and surface air temperature data were taken from the Goddard Institute for Space Studies (GISS) analysis (Hansen et al. 1999), which was supplemented in the Arctic by data from Polyakov et al. (2003). Only nearly complete time series were chosen, and missing values in the calibration period were replaced by NNR anomalies at 925 hPa after standardization. This substitution is possible because the reconstruction is based on standardized temperature anomalies and tests show high correlation between standardized temperature anomalies at the surface and 925 hPa (see Brönnimann and Luterbacher 2004).

This reconstruction technique relies on two assumptions: First, there is longer term stationarity at a given location; that is, the relationship between the surface and upper-air variables does not change on decadal scales. The second assumption is that NNR itself is a suitable basis for a reference series. This latter assumption could be problematic as NNR has known inhomogeneities (e.g., Santer et al. 1999; Randel et al. 2000). A few sample reconstructions were performed using the post-1958 radiosonde data as the upper-air series; the comparison to the NNR-based reconstruction is shown in Fig. 2. The reconstructions are virtually identical and have correlation coefficients ranging from 0.982 to 0.999; additional statistical metrics are listed in Table 1. Reconstructions based on post-IGY radiosonde data are not feasible for the majority of stations owing to short records and missing data, but the excellent agreement with NNR-based reconstructions suggests that our approach is acceptable.

c. Data preprocessing

The radiosonde data were preprocessed in several steps: The data were checked for outliers using climatology, the lapse rate was checked, and a check for basic hydrostatic consistency was performed. Because the data were launched at inconsistent and varying times throughout the record, each sounding was then adjusted to the daily mean. These daily mean adjustments were based on a diurnal cycle climatology from NNR and were calculated for each station, height, and month of the year (Brönnimann 2003b). The diurnal cycle climatology in NNR could have errors but was nonetheless

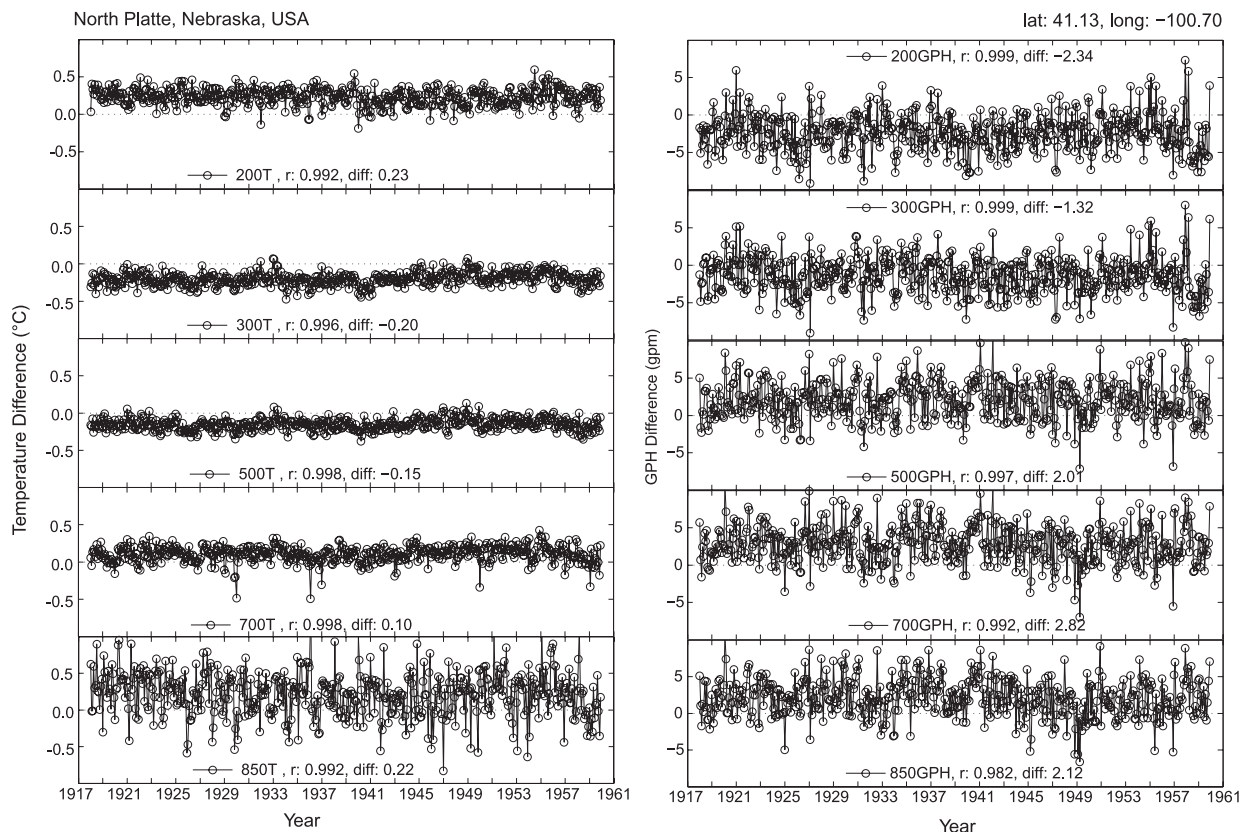


FIG. 2. Time series of the difference between the two reconstructions for North Platte, Nebraska: those based on NNR minus those based on post-1958 radiosonde data. Reconstruction differences are shown for (left) temperature and (right) GPH for 850, 700, 500, 300, and 200 hPa. Correlation coefficients and mean difference between the two reconstructions for each level and variable are given.

the most realistic choice. The diurnal cycle adjustments are very small in the free troposphere ($\sim 0.1^\circ\text{C}$ and 4 gpm) and the physically based errors that we identify are independent of this step. Additionally, calculating daily mean values for a given day at a given station using a variety of launch times resulted in daily means that were the same (the differences were not statistically significant; not shown). This suggests that the use of NNR to adjust for the launch times does not introduce a bias into the daily mean. After adjusting the soundings for launch time, they were averaged into daily and monthly means for comparison with the reference series. Monthly means were calculated only if the data met the following criteria for any given level: at least 13 soundings present in the month or no gaps longer than seven days (Brönnimann 2003b). Only temperature and geopotential height (GPH) were examined in this study.

d. Assessment basis

The data were then assessed on a station-by-station basis by comparing the monthly mean data and the monthly mean reference series at the subset of pressure

levels (850, 700, 500, 300, and 200 hPa). The difference between the two (monthly mean data – reference) is termed the bias. Levels above 200 hPa were not used due to low skill in the reconstruction (see below). For

TABLE 1. Statistics comparing the reconstructions based on NNR to those based on the data. Correlation coefficients are given for the two reconstructions (r) and for NNR vs the data for the calibration period (1960–2000, r_c) and validation period (1948–59, r_v) of the reconstruction. Differences between NNR and the data in the calibration period (NNR – data) and between the two reconstructions (Δ_{recon}) are also given.

	r	Δ_{recon}	r_v	r_c	NNR – data
850 T	0.992	0.22	0.996	0.997	0.14
700 T	0.998	0.10	0.999	0.999	0.07
500 T	0.998	-0.15	0.998	0.999	-0.14
300 T	0.996	-0.20	0.996	0.998	-0.22
200 T	0.992	0.23	0.945	0.975	0.17
850 GPH	0.982	2.12	0.993	0.975	2.19
700 GPH	0.992	2.82	0.998	0.995	2.88
500 GPH	0.997	2.01	0.999	0.998	2.26
300 GPH	0.999	-1.32	0.998	0.999	-1.84
200 GPH	0.999	-2.34	0.998	0.999	-2.87

each station a suite of diagnostic plots was created: mean bias as a function of pressure for both annual and seasonal means, time series of the bias at each level, scatterplots of the raw data at each level versus the neighboring level, and plots of the monthly mean real-valued data as a function of pressure.

As noted above, the technique employed here is based on the *a priori* assumption that there is a small number of possible errors that can be reliably identified in this manner. “Theoretical” examples of these possible errors are shown in Fig. 3, where data from a high quality station (which required no correction) has had an artificial error added to a summer and a winter profile. The error was generated by applying the correction code to an individual sounding. The radiation and lag correction follows the general framework in Väisälä (1941, 1949) and Raunio (1950) of determining the insolation at a given height, time of day, and day of year for the station and assumes a fixed lapse rate of 5 m s^{-1} (see also Brönnimann 2003b); the correction differs for each country only in the two parameters of lag-time constant and the pressure-dependence factor. This more generic correction, rather than a highly detailed instrument-specific correction, is more appropriate for the early data. The pressure correction is based on the physical principle of a measurement being recorded at the wrong pressure and is functionally similar to the lag correction.

Figure 3 shows the difference between the data with errors and the original (correct) data so as to provide a clean example of the vertical profile of these common radiosonde errors. Each panel of Fig. 3 has unique and distinct features. A units problem (geodynamic meters instead of geopotential meters, Fig. 3a) has an error only in height and is of extremely large magnitude at higher altitudes. Radiation and lag (Fig. 3b) show increasing bias with altitude for both temperature and GPH and is less than about 3°C and 100 m at the tropopause. Pressure errors (Fig. 3d) have a similar shape in the midtroposphere (both increasing with height) but have two notable differences from radiation and lag: 1) in all seasons, the temperature error drops to zero at the tropopause when the lapse rate changes sign and 2) the magnitude of the height error can reach any magnitude (and is generally substantially larger than radiation and lag). The constant temperature offset (Fig. 3c) shows a constant bias with altitude in temperature and an increasing, but small magnitude, bias in height as altitude increases (although any magnitude is possible for exceedingly large temperature offsets). We bundled radiation and lag together because, in our experience, they are either both applied or neither has been applied in an operational manner. As the lag error

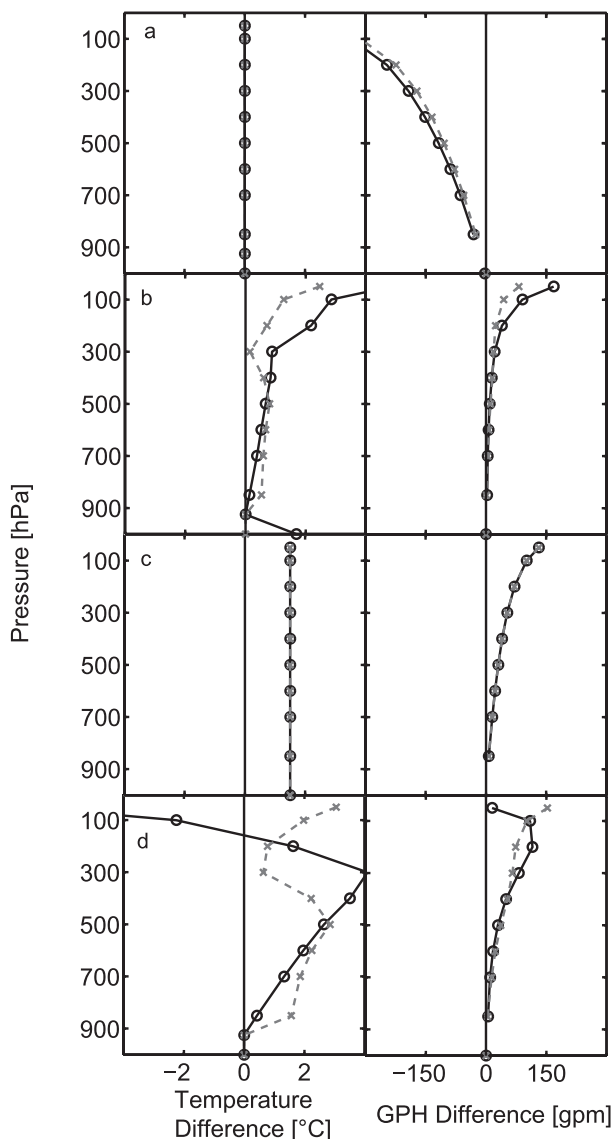


FIG. 3. Artificial errors added to a summer (black circle; 1600 UTC 12 Jul 1949) and winter (gray x; 1500 UTC 13 Jan 1951) profile from Frankfurt, Germany, for 925–50 hPa; plots show the difference between the data with the error and the original (correct) data for (a) units, (b) radiation and lag, (c) constant temperature, and (d) pressure error. Each error has a characteristic vertical profile and unique features that allow it to be distinguished from the other errors. The large error at 925 hPa in (b) the summertime radiation and lag is due to the steep gradient in the boundary layer, which causes a large lag error.

is approximately constant whereas the radiation error changes with the time of day and season, the seasonal cycle of the errors can give further information on this. Obviously, it is possible for a station to have more than one of these errors at the same time, none of them, or other errors not addressed in the current study.

e. Reconstruction skill

To assess the station quality with confidence, at least five monthly means were required and the reconstruction was required to have reasonable skill. The skill metric determined in the validation period was the reduction of error (RE) statistic (Cook et al. 1994),

$$\text{RE} = 1 - \frac{\sum_t (x_{\text{rec}} - x_{\text{obs}})^2}{\sum_t (x_{\text{null}} - x_{\text{obs}})^2},$$

where t is time, x_{rec} is the reconstructed value, x_{obs} is the observed value, and x_{null} is a null hypothesis or “no knowledge” prediction (e.g., constant, climatology, random, persistence). In our case, since we reconstruct anomalies, the chosen null hypothesis is a zero anomaly (i.e., the mean annual cycle in the calibration period): RE can range from $-\infty$ to 1, where $\text{RE} = 1$ means a perfect reconstruction, $0 < \text{RE} < 1$ means there is predictive skill in the reconstruction, and $\text{RE} = 0$ means the reconstruction is no better than the input NNR climatology (the no-knowledge prediction). A random number with the correct variance would yield an RE of -1 . Here RE is preferred over correlation or explained variance, as the latter two do not account for a bias in the reconstruction, whereas RE does.

A value of RE above 0.5 was chosen as a cutoff for reasonable skill in this work; this choice is somewhat arbitrary. An RE of 0.5 corresponds roughly to an explained variance of 50% or a correlation of 0.7. It was generally the case, however, that reconstructions with skill between 0 and 0.5 still captured climatic features reasonably well (i.e., they were still well correlated); the lower skill tended to be reflected in variability that was too small compared to NNR or the data themselves. For this reason, the low skill reconstructions were given less emphasis but were still consulted during quality assessment decisions.

f. Statistics

The target quality for bias with respect to the reference was $\pm 0.75^\circ\text{C}$ for temperature at all levels and $\pm 15\text{--}30$ gpm (increasing with height in the atmosphere and for our subset of levels) (Brönnimann 2003a). The target precision was 1.6°C and $30\text{--}80$ gpm for temperature and height (i.e., 90% of the data must lie within these limits). The targets were chosen based on the ultimate application of the postassessment data, in this case, climate variability studies; they also were chosen so as to allow identification of impact-relevant climate anomalies on a monthly-to-multiannual time scale (e.g., droughts, severe winters, monsoon changes, El Niño-related anomalies).

Statistical significance of the bias and the variance (too much variability) were calculated. The bias was determined significant if it exceeded twice the standard error, SE_{o-t} :

$$\text{SE}_{o-t} = \sqrt{\frac{s_{o-r}^2 + s_{r-t}^2}{n}}$$

in which s_{o-r}^2 is the variance of the observations with respect to the reconstructions, s_{r-t}^2 is the variance of the reconstructions with respect to the true values and is estimated from the variance of the model residuals in the calibration period, and n is the number of months with data. The variance was tested according to $s_{o-r}^2 \leq s_{\text{targ}}^2 + s_{r-t}^2$, where s_{targ}^2 is the predefined target variance (1.6°C or $30\text{--}80$ gpm) and s_{o-r}^2 was calculated from the observations and reconstructions. Because this variance is only an estimation that was sometimes based on a small sample, we used the lower 95% confidence limit of the estimated variance based on a χ^2 distribution (Brönnimann 2003b). The standard normal homogeneity test (Alexandersson and Moberg 1997) was applied to the data and the reference series, while the self-homogeneity test developed by Lanzante (1996) was applied to both time series of the raw data and of the bias. It should be noted that the statistics presented here were an aid to judgment, providing a metric for acceptable bias, but that the quality assessment process is primarily based on the vertical structure of the bias rather than minute details of the statistical tests.

g. Decisions

The final verdict for each station was based on these statistical tests combined with some expert judgment. For example, clusters of homogeneity breakpoints at the same time at different levels and for both temperature and GPH presented a stronger case for inhomogeneity than an isolated result from one of the tests. This approach was taken because this quality assessment method is primarily for judging data on a station-by-station basis (i.e., looking at the quality of the station as a whole) rather than a technique to statistically homogenize individual data points within a station time series. Initial verdicts fell into three categories: accept the station as is, correct the station, or reject the station. Stations deemed acceptable (after correction, if necessary) were further given a flag indicating higher or lower quality. Higher quality indicated the final data had nearly zero bias at all levels, while lower quality showed good agreement but with a small, nonzero bias (less than the target) in one or more levels.

To assess the quality of a station time series, the reconstruction had to have reasonably good skill, that is,

above our RE threshold of 0.5. This left a large number of initially unassessed stations where the RE was below 0.5 at all or most levels for both temperature and height. Examination of the time series of the reconstruction compared to NNR from 1948 to 1957 showed that, in the vast majority of cases, RE between 0 and 0.5 coincided with situations where the reconstruction clearly captured the climatic fluctuations in the time series but had variability that was too small. In this case, the reconstruction, although of low skill, was used in assessing the quality of the station time series. The three verdict categories (accept good quality, accept lower quality, or reject) were also applied to stations where the reconstruction skill was too low to make a definitive verdict. The poor skill should be noted and these stations should be rejected for applications requiring only very high quality data. Corrections to these stations were specified only if independent information existed (e.g., if a networkwide radiation and lag correction was relevant) but no constant temperature adjustments were made. A small number of stations were totally unassessed owing to insufficient data for assessment (less than five monthly means) or insufficient input data to reconstruct the reference series.

h. Corrections

With this technique, the diagnosis of problems is separate from the correction in most cases. Rather than performing statistical homogeneity adjustments, the majority of corrections were performed if the bias was characteristic of a problem with a known physical basis, such as uncorrected or undercorrected radiation and lag errors, pressure sensor errors, temperature sensor offsets, or units problems (e.g., geodynamic rather than geopotential meters). The corrections involved three stages: 1) the errors were preliminarily diagnosed via their characteristic vertical profile of the bias (see Fig. 3); 2) the corrections were applied based on independent information about the data [e.g., radiation correction as given in Teweles and Finger (1960)] so that independence from the NNR-based reconstructions is maintained (and thus independence from NNR and its inhomogeneities); and 3) checking whether the application of the correction removed the bias in both temperature and GPH. The third criterion was critical—in some cases, the original bias appeared to be of one type but the correction failed to resolve it. The failure in some cases meant a new bias appeared (e.g., temperature was “fixed” but then GPH had a bias) or there was a remaining bias with an irregular profile. The only correction that used information from the reference series and was thus dependent on NNR (via the reconstructions) was the constant temperature correction.

For many applications the correction to NNR as a standard is desirable, and it is consistent within our framework (e.g., the use of NNR climatologies to adjust the diurnal cycle and to present anomalies). For other applications it may be more desirable to correct the series with respect to another standard.

Although Teweles and Finger (1960) refer to “solar radiation temperature correction *used* by various countries” (emphasis added), data from the former Soviet Union all showed signs of uncorrected radiation and lag (RL) errors, causing some confusion about whether the published corrections had been applied operationally. Nagurny (1998) also applied RL corrections to all former Soviet Union data before 1957, confirming the diagnosis of uncorrected RL errors. The specific ending date for the RL corrections was determined for each station from the diagnostic plots. After correction, some stations still had significant upper-level warm anomalies up to and including 1954, suggesting that the earlier instruments might require an even stronger correction than that published in Teweles and Finger (1960), although the exact value of that larger correction remains undetermined. In general, it seems that the use of country to identify radiosonde type is an acceptable approach for the early data for two reasons: during this time, the radiosonde was generally being developed at a governmental level for national or regional use rather than as a commercial product, and the radiation and lag correction that we use follows a generic framework (see section 3d). The missing metainformation clearly is a problem and the corrections have to be reassessed once more information becomes available.

In all cases, decisions and correction amounts were based on the aggregate monthly mean but were applied to individual soundings, which were reprocessed into new monthly means. Corrections were made to the temperature and then geopotential height was recalculated using the hydrostatic equation (Brönnimann 2003b).

4. Example stations

The mean bias of both temperature and height as a function of pressure is the centerpiece of the quality assessment technique, summarizing the overall quality of the station. Figure 4 shows four example stations with correction type, Fig. 4a: accepted as is (Caribou, Maine), Fig. 4b: radiation and lag correction (Dikson Island, Russia), Fig. 4c: constant temperature offset (Wernigerode, Germany), and Fig. 4d: rejected due to large and unidentifiable errors (Warsaw, Poland). For the two stations that were adjusted (Figs. 4b,c), both

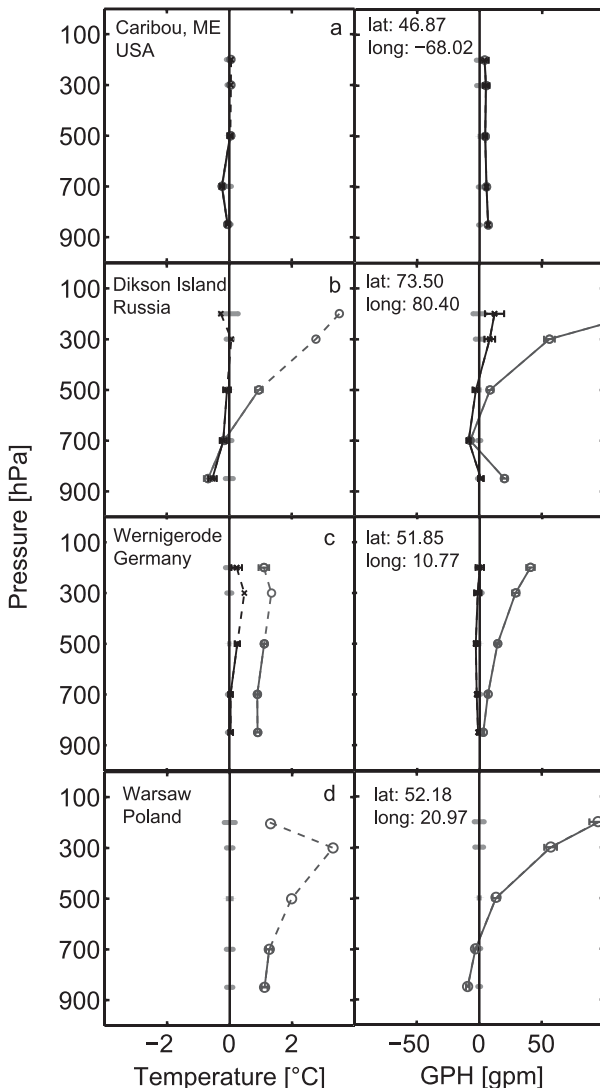


FIG. 4. Bias of the data relative to the reference as a function of pressure for (left) temperature and (right) GPH; data are shown for the beginning of the record to the end of 1957 or the end of the record if before 1957. Dark gray lines represent before correction and black lines after correction. Dashed lines indicate that the reconstruction did not meet the predefined metric in quality. Error bars about the data are $\pm 1 SE_{o-r}$ and those about the zero line (light gray) are $\pm 1 SE_{r-t}$ of the reconstruction. Examples are shown for a station requiring (a) no correction (Caribou, Maine), (b) radiation and lag correction (Diskon Island, Russia), (c) constant temperature adjustment (Wernigerode, Germany), and (d) a rejected station (Warsaw, Poland). Note that there is no corrected (black) line shown for Warsaw as it was rejected.

before and after are shown. The error bars about the data are $\pm 1 SE_{o-r}$ (of the data) and the error bars around the zero line are $\pm 1 SE_{r-t}$ (of the reconstruction).

The skill of the reference series was low for some stations and variables. Levels with good skill are shown as circles (before correction) and stars (after correction)

with heavy, solid lines. The poor skill levels are plotted with a dashed line; low skill levels were given no or very little weight in determining the quality of the data, depending on the availability of other, high skill levels. However, they can be useful in determining the overall consistency of a correction, especially those that were applied networkwide.

The typical data problems have characteristic vertical structure in the bias as seen in Fig. 3. For Caribou, Maine (Fig. 4a), the station accepted “as is,” the bias is very close to zero at all levels for both variables. In the radiation and lag example (Diskon Island, Russia, Fig. 4b; cf. Fig. 3b), the temperature and GPH biases increase with height as shown by the gray line; after the radiation and lag correction (black lines with stars), the data fall into line with the reconstruction. Constant temperature offsets were specified for stations that show the same temperature offset at all levels: a constant temperature bias with height and a slightly increasing GPH bias with height (e.g., Wernigerode, Germany, Fig. 4c; cf. Fig. 3c). In most cases, it is unclear if the data are biased or the reconstruction is; this adjustment reflects the fact that the apparent disagreement between the data and the reconstruction is resolved with a constant temperature adjustment. For this reason they are termed adjustments rather than corrections, as they adjust the radiosonde data to the NNR 1960–2000 climatology used in the reconstructions. The adjustment was considered mandatory, if the offset is both significant ($>2 SE_{o-r}$) and larger than the target, or optional if the offset was significant but less than the target.

Some stations were found to have unidentifiable and uncorrectable errors: for example, positive temperature bias and negative GPH bias (or vice versa) or biases that changed erratically over time. In these cases, correcting one variable would make the other worse for part or all of the record (e.g., Warsaw, Poland, Fig. 4d). If these uncorrected errors were larger than the target ($\pm 0.75^\circ\text{C}$ and $\pm 15\text{--}30\text{ gpm}$), the station was rejected; if they were within these limits, the station was accepted but coded as lower quality. In the case of Warsaw, the large error in the bias plot is clearly reminiscent of a pressure sensor error (cf. Fig. 3d). However, the conflicting temperature and height errors (positive temperature bias with negative height bias at low levels) and a worsening of some levels and variables after applying the proposed correction led to a decision to reject the entire station. This decision was also made because the pressure correction requires good skill in the upper troposphere to fine-tune the magnitude of the pressure error (an iterative process); the low skill at higher altitudes precludes this adjustment at this time.

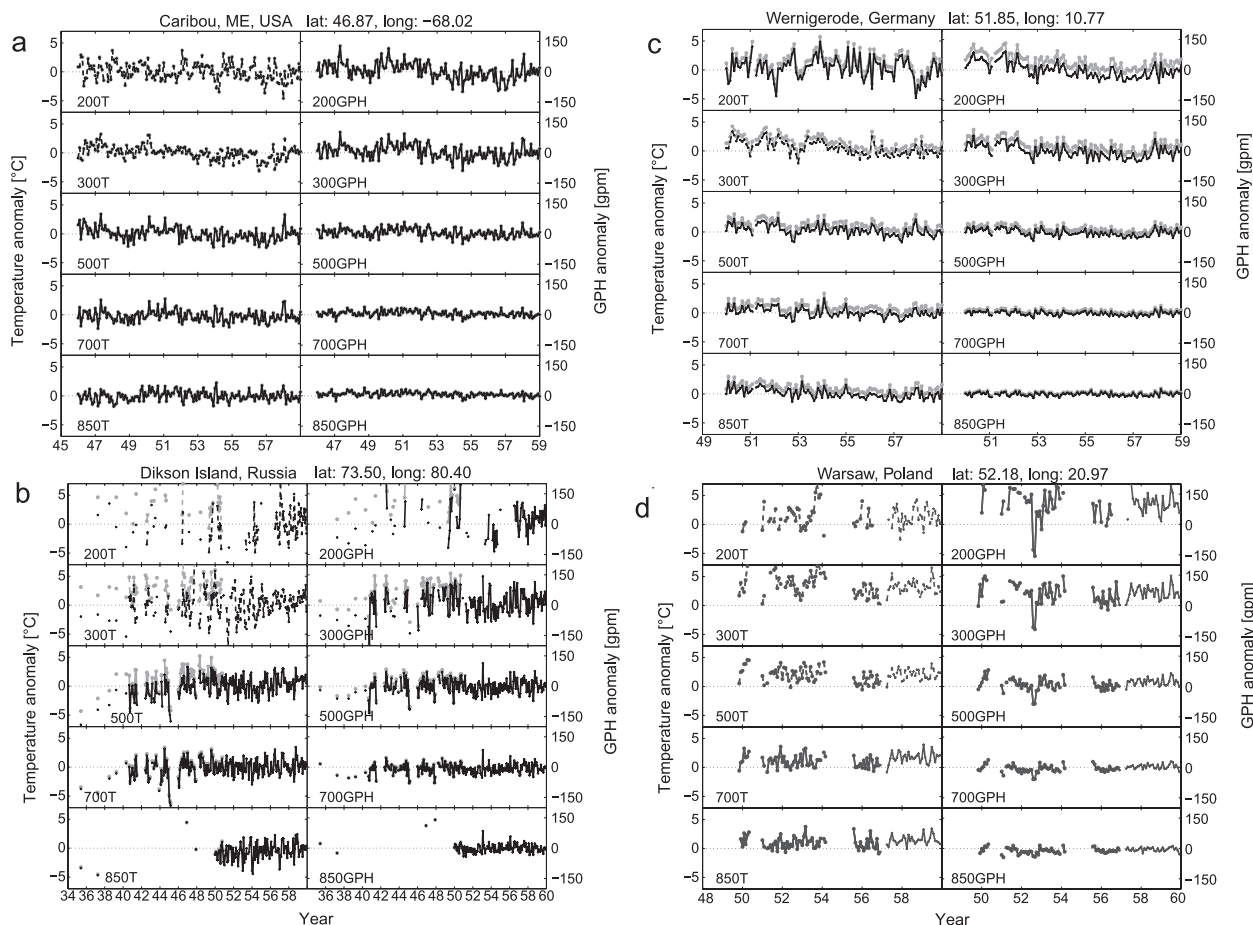


FIG. 5. Time series of temperature (left within pair) and height (right within pair) relative to the reconstruction at 850, 700, 500, 300, and 200 hPa. Gray lines show before correction and black lines show after correction. Dashed lines indicate that the reconstruction did not meet statistical significance in quality. Examples are shown for the corresponding stations as in Fig. 4.

About one-third of the stations could not be assessed with complete confidence owing to poor skill in the reconstruction (below 0.5 at most or all levels in both temperature and GPH). Because the reconstruction does still capture some information about the climate, the vertical bias plots were examined and the unassessed stations were categorized as accept or reject, although the lack of reconstruction skill should be kept in mind. Some of these stations with poor skill were specified as needing a radiation and lag correction if they were part of a networkwide correction, such as in the former Soviet Union, and the low-skill reconstruction plots showed improvement after the correction, confirming their usefulness. No constant temperature adjustments were specified for these stations.

In Fig. 5, time series for the examples shown in Fig. 4 are presented for both temperature and height on the subset of pressure levels used for quality assessment. In the case of the rejected station, the decision was based

on the large and conflicting errors in temperature and pressure. In particular, 700 hPa has a large positive temperature bias but a small positive GPH bias alternating with no GPH bias, whereas 500 hPa shows a positive temperature bias but nearly zero GPH bias and 200 hPa shows a small temperature bias and a large positive GPH bias except for some months in 1952. The poor skill in the temperature reconstruction complicates the picture. The accepted stations, on the other hand, show biases that are internally physically consistent and relatively stable over time.

Figure 6 shows a time series of bias at 500 hPa for Yekaterinburg (formerly Sverdlovsk), Russia, a station with a units problem. When upper-air measurements were first developed, heights were normally reported in geodynamic meters rather than geopotential meters. The difference between geodynamic and geopotential meters increases with height and can be hundreds of meters in the stratosphere yet is very difficult to detect

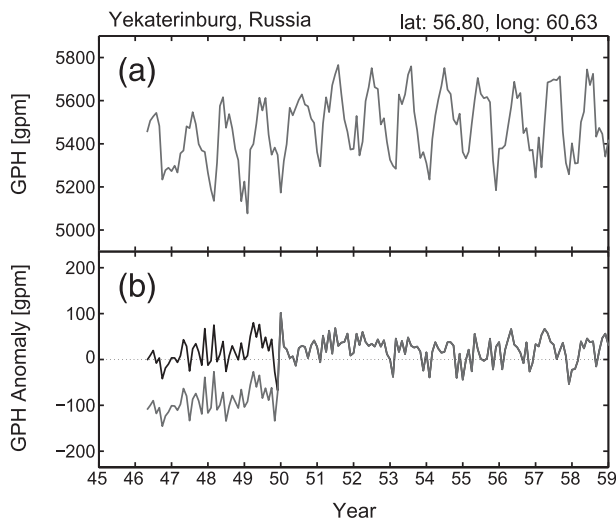


FIG. 6. Time series of (a) raw data and (b) anomaly from the reconstruction for Yekaterinburg (formerly Sverdlovsk), Russia, at 500 hPa, before (gray) and after (black) units correction.

in time series of real-valued data, such as in Fig. 6a. In a time series of anomalies from the reference (Fig. 6b), the problem is easily detected.

Time series of the number of corrected stations (as a fraction of the number of stations) and the total number of stations are shown in Fig. 7. Note that, in the time series of corrected stations, individual stations may be represented more than once at any time step if they had multiple corrections and that the number of stations is presented on a log scale. The sparsity of data in the early years causes the correction time series to appear jumpy before 1940. In some cases the ending time for the correction was clear (e.g., the networkwide radiation and lag correction in the former Soviet Union and the units problems for which there was a clear step-function change in the bias time series). In other cases the correction was applied to the entire time period; this is especially the case with the constant temperature adjustments. Because this study focused on pre-IGY data, this led to an unspecified ending time for many corrections. Corrections were identified and applied only up to the end of 1957 (or earlier, if specified).

These examples cover simple, easily corrected errors in the data that have nonetheless eluded detection using standard quality control procedures, but that we were able to identify by examining anomalies from a reconstructed reference series. A map showing station data before and after correction, differences, and required correction type is shown in Fig. 8, revealing how widespread these errors are. Units and radiation and lag discrepancies were generally confined to the former Soviet Union, whereas the constant temperature corrections

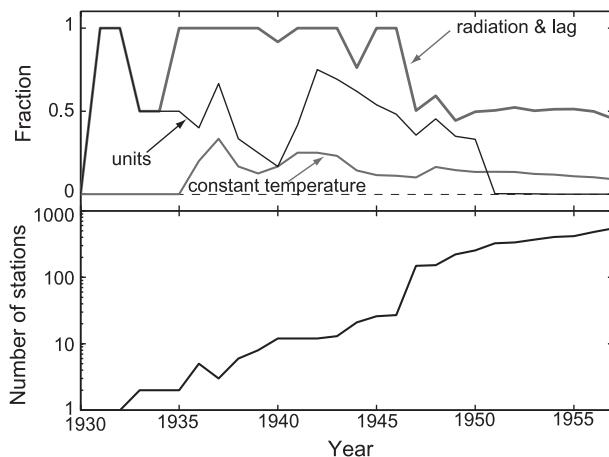


FIG. 7. Number of stations (top) that required a correction as a fraction of the number of stations with data for each year and (bottom) the total number of stations.

reveal no underlying pattern. This is discussed in more detail in section 5.

During the cross-check of duplicated stations in the original data collected for this study, some stations were identified with incorrect timestamps. In one example, Aktjubinsk, Kazakhstan, the data from the TD-54 source had two soundings per day (0300 and 1500 UTC), while the same station in the IGRA database had four soundings per day (0300, 0500, 1500, and 1700 UTC).

On closer inspection, the two additional soundings in IGRA (0500 and 1700 UTC) were found to be duplicates of the other two in IGRA but with a timestamp two hours later, consistent with the station being at UTC + 3. A complete list of such stations was not made, as the decision to prioritize other sources over IGRA had already been made and the problem was found only in the early part of the record. Where noted, this information is added to the cross-referenced station list. Table 2 presents a summary of statistics describing the correction types both globally and by network.

5. Large-scale impacts of the corrections

The widespread geographical extent of the corrections has implications for hemispheric-scale analysis figures because the extent of the errors is on the same scale as large-scale atmospheric features such as planetary waves. Station data for 1956 showing 500-hPa temperature and height are presented in Fig. 8. The original and corrected data are shown as well as the difference between them; data are shown as anomalies from a 1961–90 NNR climatology for each station and the correction type is coded by the symbol in the difference plots. The corrected data are colder and the

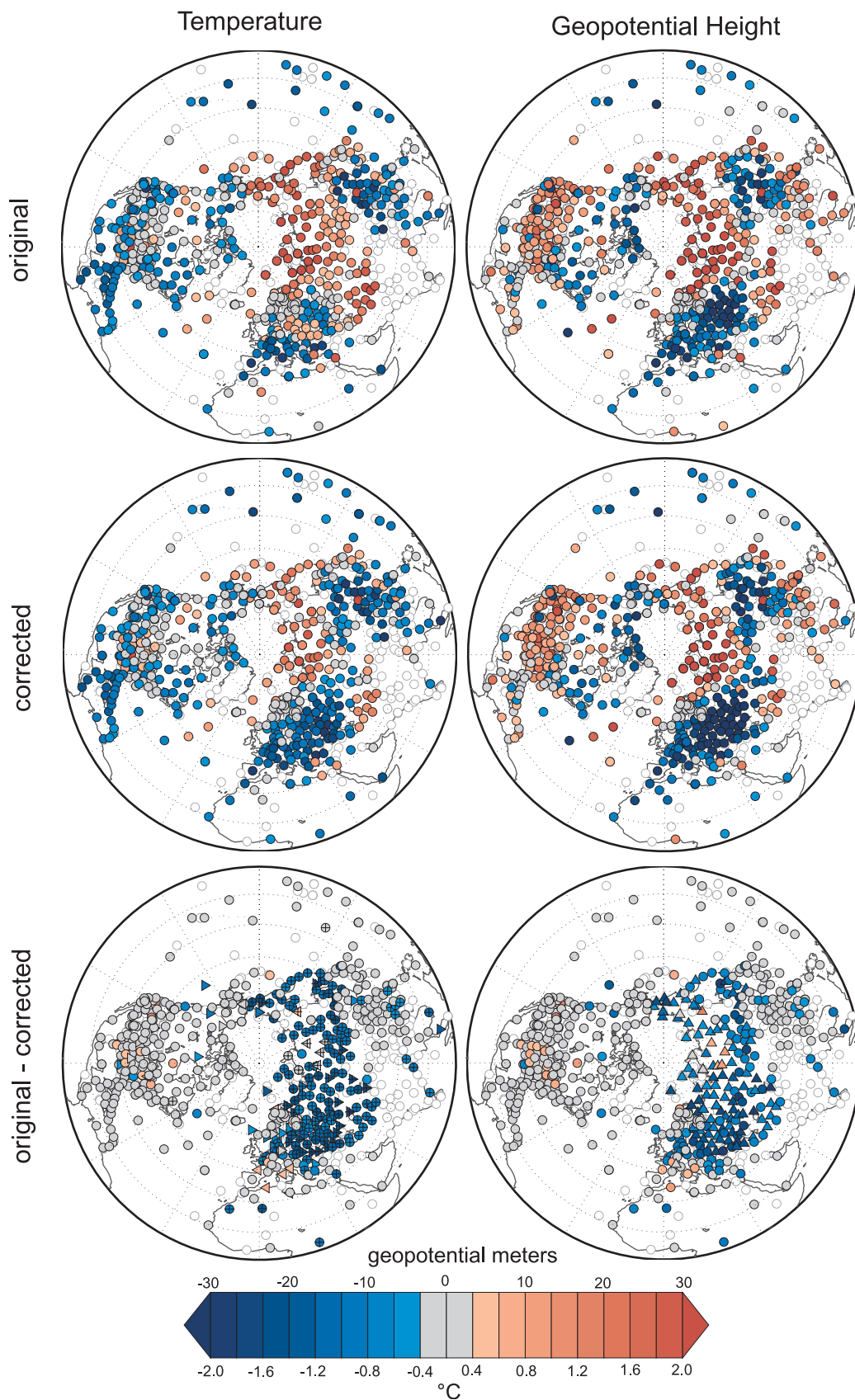


TABLE 2. Stations requiring correction by sonde type as a proxy for network (see station list for complete details). The former Soviet Union stations are labeled SU: “SU strong” for the stations requiring a stronger-than-published correction up to and including 1954.

	Vaisala	U.S.	SU	SU strong	Total
Radiation and lag	31	3	177	26	237
GPH units	3	0	64	18	85
Constant T (bias: +, -)	29 (18,11)	23 (15,8)	39 (32,7)	5 (4,1)	96 (69,27)
Optional const T (bias: +, -)	7 (7,0)	42 (7,35)	45 (45,0)	4 (4,0)	98 (63,35)
No correction	106	279	55	0	440
Accepted, high quality	109	296	141	26	572
Accepted, lower quality	66	50	99	0	215
Rejected	24	9	28	0	61
Unassessed (no data or recon)	20	11	0	0	31
Total	219	366	268	26	879

heights are lower across the whole of Eurasia, reflecting the networkwide radiation and lag corrections. Constant temperature adjustments are more isolated and do not exhibit such coherency in sign and scale. The temperature and height fields match each other better, and thus are more physically consistent, after correction. The spatial structure of the correction is important, especially for spatially large networks such as the former Soviet Union, which introduce errors of the spatial scale of planetary waves. The unit corrections (necessary only for data prior to 1950 in most cases) caused a raising of the layer that is larger than the magnitude of the radiation and lag correction, leading to a situation in which temperatures decreased but heights increased after correction before 1950.

A considerable fraction of the data assessed in this study was assimilated into NNR. Examination of real-valued annual mean data compared to NNR reveals that NNR has a warm bias over much of Eurasia in the 1950s. Figure 9 shows annual mean NNR minus annual-mean station data in 1956 for the Northern Hemisphere using both the original and the corrected data. NNR has a very slight cold bias over Eurasia compared to the uncorrected data, but the errors are less coherent in sign and location and suggest that NNR is well constrained by the assimilated (erroneous) data. When compared against the corrected data, however, NNR has a coherent warm and high bias over Eurasia. This hemispheric-scale bias would have an impact on analyses of NNR prior to 1958, particularly on principal component analyses of large-scale fields such as teleconnection studies. NNR

prior to 1958 should be treated with some caution. Note that radiosondes are not the only source of error in NNR and that, conversely, the new radiosonde product is by no means error free. For example, the topography of the Tibetan Plateau is clearly revealed by the large swath of negative height biases in both before and after figures. It is well known that NNR has some difficulty over significant terrain features, in particular showing a cold and low bias over the Himalayas (e.g., Xie et al. 2008). The reconstructions in this area consistently had low skill, hampering our ability to stringently assess the quality of stations in this region.

6. Conclusions

Worldwide radiosonde data from prior to 1958 were compiled and reevaluated using a novel quality assessment technique that creates a statistically reconstructed reference series. Errors that otherwise prove difficult to detect are identified more clearly when the data are compared to such a reference series. Pervasive errors were found in some networks over the Eurasian continent affecting both temperature and height. In particular, radiation and lag errors were common; these errors, which affect both temperature and height, are easily corrected using published corrections (Teweles and Finger 1960). Unit errors in the geopotential height were also common. These simple errors have a large impact on resulting analyses of the spatial field as they affect data over a geographical area that is on the same

←

FIG. 8. Annual mean anomaly of 500-hPa (left) temperature and (right) GPH for 1956 from (top) original data, (middle) corrected data, and (bottom) difference (corrected – original). Anomalies are from a 1961–90 NNR climatology. In the difference temperature (lower left), radiation and lag corrections are indicated with a plus and constant temperature corrections are indicated with right and left facing triangles for positive and negative biases, respectively. White circles indicate a station with no (or rejected) data. In the GPH difference (lower right), units corrections are indicated by upward pointing triangles. The radiation and lag and constant temperature corrections affect both temperature and height, but the symbols are separated for clarity.

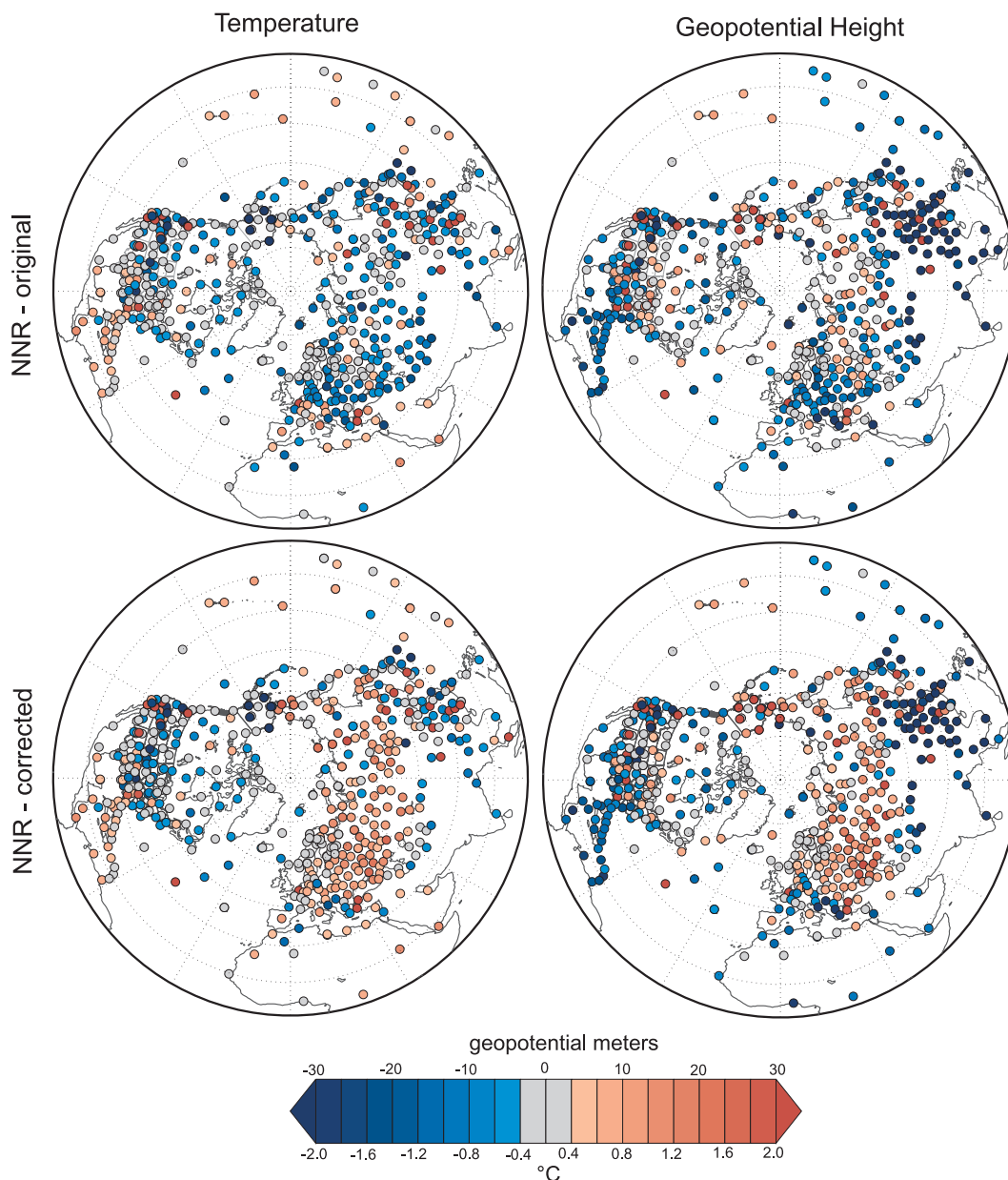


FIG. 9. 1956 Annual mean NNR minus (top) raw and (bottom) corrected (left) temperature and (right) GPH at 500 hPa. NNR appears to have a very slight cool bias relative to the uncorrected data but generally shows good agreement with some random errors. NNR has a widespread warm and high bias relative to the corrected data owing to the assimilation of uncorrected data. Note the persistent low bias over the Tibetan Plateau where NNR has difficulty with the terrain.

scale as atmospheric phenomena, such as planetary waves that can be analyzed through modes of variability (such as the Arctic and North Atlantic Oscillations or the Pacific–North America pattern). Some stations were also found to have a constant temperature offset relative to the reconstruction that must be corrected if the data are to be merged with NNR for other purposes. The corrected data show improved internal consistency be-

tween temperature and height. Of 879 unique stations, 787 were of sufficient quality after quality assessment and correction (if necessary) to be kept, while 92 stations were rejected owing to unresolvable errors or data sparsity. [Monthly-mean data are available in an online archive at <http://www.historicalupperair.org>, along with a complete index of stations (cross-referenced to several source archives) and the verdicts.]

Examination of NNR revealed that the reanalysis closely follows the assimilated data, leading to a widespread warm and high bias in the 1950s due to assimilation of large amounts of uncorrected data. Caution should be used in interpreting hemispheric-scale analysis of NNR prior to 1958.

The homogenization of historical radiosonde data is a difficult process that has been tackled many times. Although errors undoubtedly remain in the data, the corrections presented here are one step in the process of increasing the usefulness of the earlier data and providing the climate community with a new comprehensive dataset.

Acknowledgments. We thank the anonymous reviewers for helpful comments, which resulted in a much-improved paper. This work was supported by the Swiss National Science Foundation. The TD54 and CARDS542 datasets were kindly provided by Joey Comeaux at NCAR, and the TD6201 dataset by Joe Elms at NCDC. The Lindenberg, Germany, data were provided by the German Meteorological Service, the Payerne data by MeteoSwiss, and the Ilmala data by the Finnish Meteorological Institute.

REFERENCES

- Alexandersson, H., and A. Moberg, 1997: Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *Int. J. Climatol.*, **17**, 25–34.
- Allan, R., and T. Ansell, 2006: A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *J. Climate*, **19**, 5816–5842.
- Brönnimann, S., 2003a: Description of the 1939–1944 upper air data set (UA39_44) version 1.0. Lunar and Planetary Laboratory, University of Arizona Tech. Rep., 41 pp.
- , 2003b: A historical upper air-data set for the 1939–44 period. *Int. J. Climatol.*, **23**, 769–791.
- , and J. Luterbacher, 2004: Reconstructing Northern Hemisphere upper-level fields during World War II. *Climate Dyn.*, **22**, 499–510.
- , C. Mohr, T. Ewen, and A. Grant, 2005: The Finnish radiosonde records from Ilmala, Rovaniemi, and Äänilinna, 1942–1947. ETH Zurich Tech. Rep., 13 pp.
- Cook, E., K. Briffa, and P. Jones, 1994: Spatial regression methods in dendroclimatology: A review and comparison of two techniques. *Int. J. Climatol.*, **14**, 379–402.
- Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the Integrated Global Radiosonde Archive. *J. Climate*, **19**, 53–68.
- Eskridge, R., O. Alduchov, I. Chernykh, Z. Panmao, A. Polansky, and S. Doty, 1995: A Comprehensive Aerological Reference Data Set (CARDS): Rough and systematic errors. *Bull. Amer. Meteor. Soc.*, **76**, 1759–1775.
- , J. Luers, and C. Redder, 2003: Unexplained discontinuity in the U.S. radiosonde temperature data. Part I: Troposphere. *J. Climate*, **16**, 2385–2395.
- Ewen, T., A. Grant, and S. Brönnimann, 2008: A monthly upper-air data set for North America back to 1922 from the *Monthly Weather Review*. *Mon. Wea. Rev.*, **136**, 1792–1805.
- Free, M., and Coauthors, 2002: Creating climate reference datasets: CARDS workshop on adjusting radiosonde temperature data for climate monitoring. *Bull. Amer. Meteor. Soc.*, **83**, 891–899.
- , D. J. Seidel, J. Angell, J. K. Lanzante, I. Durre, and T. C. Peterson, 2005: Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new data set of large-area anomaly time series. *J. Geophys. Res.*, **110**, D22101, doi:10.1029/2005JD006169.
- Fu, Q., C. Johanson, S. Warren, and D. Seidel, 2004: Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429**, 55–58.
- Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate*, **20**, 1377–1403.
- Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, 1999: GISS analysis of surface temperature change. *J. Geophys. Res.*, **104**, 30 997–31 022.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.
- Lanzante, J., 1996: Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.*, **16**, 1197–1226.
- , S. Klein, and D. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *J. Climate*, **16**, 224–240.
- Mears, C., and F. Wentz, 2005: The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science*, **309**, 1548–1551.
- Nagurny, A., 1998: Climatic characteristics of the tropopause over the Arctic basin. *Ann. Geophys.*, **16**, 110–115.
- Parker, D. E., M. Gordon, D. P. M. Cullum, D. M. H. Sexton, C. K. Folland, and N. Rayner, 1997: A new global gridded radiosonde temperature data base and recent temperature trends. *Geophys. Res. Lett.*, **24**, 1499–1502.
- Polyakov, I., R. Bekryaev, G. Alekseev, U. Bhatt, R. Colony, M. Johnson, A. Maskhtas, and D. Walsh, 2003: Variability and trends of air temperature and pressure in the maritime Arctic, 1875–2000. *J. Climate*, **16**, 2067–2077.
- Randel, W., F. Wu, and D. Gaffen, 2000: Interannual variability of the tropical tropopause derived from radiosonde data and NCEP reanalyses. *J. Geophys. Res.*, **105**, 15 509–15 523.
- Raunio, N., 1950: Amendments to the computation of the radiation error of the Finnish (Väisälä) radiosonde. *Geophysica*, **4**, 14–20.
- Santer, B. D., J. Hnilo, T. Wigley, J. Boyle, C. Doutriaux, M. Fiorino, D. Parker, and K. Taylor, 1999: Uncertainties in observationally based estimates of temperature change in the free atmosphere. *J. Geophys. Res.*, **104**, 6305–6333.
- , and Coauthors, 2005: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science*, **309**, 1551–1556.
- Sherwood, S., J. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late-20th century warming. *Science*, **309**, 1556–1559.

- Teweles, S., and F. Finger, 1960: Reduction of diurnal variation in the reported temperatures and heights of stratospheric constant-pressure surfaces. *J. Meteor.*, **17**, 177–194.
- Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res.*, **110**, D18105, doi:10.1029/2004JD005753.
- Väisälä, V., 1941: Der Strahlungsfehler der finnischen Radiosonde. *Mitteilungen des Meteorologischen Instituts der Universität Helsinki* 47, 62 pp.
- , 1949: Solar radiation intensity at the ascending radiosonde. *Geophysica*, **3**, 37–55.
- Xie, A., J. Ren, X. Qin, and S. Kang, 2008: Pressure and temperature feasibility of NCEP/NCAR reanalysis data at Mt. Everest. *J. Mt. Sci.*, **5**, 32–37.